575

NPS-AS-93-028

NAVAL POSTGRADUATE SCHOOL
Monterey, California

CRAMÉR-VON MISES VARIANCE ESTIMATORS
FOR SIMULATIONS

by

David Goldsman
Keebom Kang
Andrew F. Seila

September 1993

Approved for public release; distribution unlimited.

Prepared for: Naval Postgraduate School
Monterey, CA 93943-5000

NAVAL POSTGRADUATE SCHOOL
Monterey, CA 93943-5000

Rear Admiral Thomas A. Mercer
Superintendent

This report was prepared for Naval Postgraduate School and funded by O&MN direct funding.

This report was prepared by:

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave Blank)*

2. REPORT DATE
September 1993

3. REPORT TYPE AND DATES COVERED
Technical Report

4. TITLE AND SUBTITLE
Cramér-von Mises Variance Estimators for Simulations

5. FUNDING NUMBERS
O&MN direct funding

6. AUTHOR(S)
Goldsman, David; Kang, Keebom; Seila, Andrew F.

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER
NPS-AS-93-028

9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

We study estimators for the variance parameter $\sigma^2$of a stationary process. The estimators are based on weightings yield estimators that are "first-order unbiased" for σ^2. We derive an expression for the asymptotic variance of the new estimators; this expression is then used to obtain the first-order unbiased estimator having the smallest variance among fixed-degree polynomial weighting functions. Although our work is based on asymptotic theory, we present exact and empirical examples to demonstate the new estimators' small-sample robustness.

14. SUBJECT TERMS
Simulation, stationary process, variance estimation, standardized time series, cramer-von mises estimator

15. NUMBER OF PAGES
16

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassifed	Unclassifed	Unclassifed	

NSN 7540-01-280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std 23908
298-102

Cramér-von Mises Variance Estimators for Simulations

David Goldsman Keebom Kang Andrew F. Seila

Abstract

We study estimators for the variance parameter σ^2 of a stationary process. The estimators are based on weighted Cramér-von Mises statistics, and certain weightings yield estimators that are "first-order unbiased" for σ^2. We derive an expression for the asymptotic variance of the new estimators; this expression is then used to obtain the first-order unbiased estimator having the smallest variance among fixed-degree polynomial weighting functions. Although our work is based on asymptotic theory, we present exact and empirical examples to demonstrate the new estimators' small-sample robustness.

Authors' addresses: David Goldsman, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332, sman@isye.gatech.edu; Keebom Kang, Department of Administrative Sciences, Naval Postgraduate School, Monterey, CA 93943, 5030p@navpgs.bitnet; Andrew F. Seila, Terry College of Business, University of Georgia, Athens, GA 30602-6255, aseila@cbacc.cba.uga.edu.

Keywords: Simulation, Stationary Process, Variance Estimation, Standardized Time Series, Cramér-von Mises Estimator.

1 Introduction

Consider a stationary process $Y_1, Y_2, \ldots, Y_n$ with mean μ. Such processes are often encountered in the context of steady-state simulation. For instance, the Y_i's might represent successive customer transit times in a complicated queueing system that has been run to steady state. If one is interested in estimating μ, the obvious unbiased point estimator is the sample mean $\overline{Y}_n$. A measure of the sample mean's precision is $\mathsf{Var}(\overline{Y}_n)$, which is unknown.

In this article, we investigate new estimators for $\mathsf{Var}(\overline{Y}_n)$, or equivalently, for $\sigma_n^2 \equiv n\,\mathsf{Var}(\overline{Y}_n)$. A related quantity is also of interest—the *variance parameter*, $\sigma^2 \equiv \lim_{n\to\infty} \sigma_n^2$. The literature studies many variance estimation methods; e.g., batch means, independent replications, spectral analysis, overlapping batch means, regeneration, autoregressive modeling, and standardized time series (STS) (see [3]). The estimators presented herein are based on weighted functionals of standardized time series. We shall show that the new estimators are asymptotically unbiased for σ^2 and that they have lower variance than competing estimators.

We first give some background. The *standardized time series* is defined as

$$T_n(t) \equiv \frac{\lfloor nt \rfloor(\overline{Y}_n - \overline{Y}_{\lfloor nt \rfloor})}{\sigma\sqrt{n}} \text{ for } 0 \le t \le 1,$$

where $\overline{Y}_j \equiv \sum_{k=1}^{j} Y_k/j$, $j = 1,\ldots,n$, and $\lfloor\cdot\rfloor$ is the greatest integer function (Schruben [20]). Under mild conditions (see Foley and Goldsman [8], Glynn and Iglehart [9], or Schruben [20]), one can show that

$$(\sqrt{n}(\overline{Y}_n - \mu), \sigma T_n) \Rightarrow (\sigma\mathcal{W}(1), \sigma\mathcal{B}),$$

where $\mathcal{W}$ is a standard Brownian motion process, $\mathcal{B}$ is a standard Brownian bridge process on $[0,1]$, and $\Rightarrow$ denotes weak convergence (as n becomes large) on $D[0,1]$, the space of right-continuous functions on $[0,1]$ having left-hand limits. It is well-known that all finite-dimensional joint distributions of $\mathcal{B}$ are normal and $\mathsf{Cov}(\mathcal{B}(s),\mathcal{B}(t)) = \min(s,t)(1-\max(s,t))$, $0 < s,t < 1$. Further, one can express $\mathcal{B}(t) = t\mathcal{W}(1) - \mathcal{W}(t)$; so it is easy to show that $\mathcal{W}(1)$ and $\mathcal{B}$ are independent, and thus $\sqrt{n}(\overline{Y}_n - \mu)$ and σT_n are asymptotically independent.

The remainder of the paper is organized as follows. §2 reviews the STS weighted area estimator for σ^2; the weighted area estimator will serve as a benchmark for comparison in the subsequent sections. §3 presents new estimators similar to weighted Cramér-von Mises (CvM) statistics, and establishes some of their properties. In particular, we find a class of CvM estimators that is "first-order unbiased" for σ^2; these estimators also have lower variance than that of the weighted area estimator. Performance of the estimators is studied in §§4 and 5, where we present exact and empirical results, respectively. §6 summarizes and discusses the results of the previous sections. §7 proposes a number of augmentations to the basic CvM estimator and provides conclusions.

2 The Weighted Area Estimator

We start with a discussion of the so-called weighted area estimator for σ^2, first popularized by Schruben [20]. Suppose we define

$$A(f;n) \equiv \frac{\sum_{k=1}^{n} f(\frac{k}{n})\sigma T_n(\frac{k}{n})}{n}$$

and

$$A(f) \equiv \int_0^1 f(t)\sigma\mathcal{B}(t)\,dt,$$

where $f(t)$ is continuous on $[0,1]$, not dependent on n, not identically zero, and normalized so that $\mathsf{Var}(A(f)) = \sigma^2$, i.e.,

$$\mathsf{Var}\left(\int_0^1 f(t)\mathcal{B}(t)\,dt\right) = 2\int_0^1\int_0^t f(s)f(t)s(1-t)\,ds\,dt = 1. \tag{1}$$

(The above expression can be simplified a great deal; see [12].) Thus, $A(f) \sim \sigma\mathsf{Nor}(0,1)$. Goldsman and Schruben [13] (also see [7], [9], and [20]) show that $A(f;n) \xrightarrow{\mathcal{D}} A(f)$, where $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution as n becomes large. So by the continuous mapping theorem (see Theorem 5.1 of Billingsley [2]),

$$A^2(f;n) \xrightarrow{\mathcal{D}} A^2(f) \sim \sigma^2\chi_1^2.$$

Since the limiting random variable $A(f)$ is the weighted area under a Brownian bridge process, we refer to $A^2(f;n)$ as the *weighted area* estimator for σ^2.

Before stating the main result on the weighted area estimator, we introduce notation that will be useful in the sequel. Define the covariance function $R_k \equiv \mathsf{Cov}(Y_1, Y_{1+k})$, $k = 0, \pm1, \pm2, \ldots$, and the quantities $\gamma \equiv -2\sum_{k=1}^{\infty} kR_k$, $F \equiv \int_0^1 f(s)\,ds$, and $\overline{F} \equiv \int_0^1\int_0^t f(s)\,ds\,dt$. Further, the notation $p(n) = O(q(n))$ means that $|p(n)/q(n)| \le K$ as $n \to \infty$ for some constant K, and $p(n) = o(q(n))$ means that $p(n)/q(n) \to 0$ as $n \to \infty$.

The main theorem for the weighted area estimator is as follows.

Theorem 1 (see [8] and [12]) Under mild conditions on the covariance function, the expected value of the weighted area estimator is

$$\mathsf{E}[A^2(f;n)] = \sigma^2 + \frac{[(F-\overline{F})^2 + \overline{F}^2]\gamma}{2n} + o(1/n).$$

Under an additional uniform integrability assumption (see Billingsley [2]'s Theorem 5.4 and its preceding comments), the asymptotic variance of the weighted area estimator is $\mathsf{Var}(A^2(f)) = \mathsf{Var}(\sigma^2\chi_1^2) = 2\sigma^4$.

Example 1 The expected value of the weighted area estimator with constant weighting function $f_0(t) \equiv \sqrt{12}$, for all $t \in [0,1]$, is $\mathsf{E}[A^2(f_0;n)] = \sigma^2 + 3\gamma/n + o(1/n)$. □

Henceforth, if the bias of an estimator for some parameter is $o(1/n)$, we shall say that the estimator is *first-order unbiased* for that parameter.

Example 2 If the weighting function $f(t)$ satisfies $F = \overline{F} = 0$ (in addition to the normalizing condition (1)), then Theorem 1 says that $\mathsf{E}[A^2(f;n)] = \sigma^2 + o(1/n)$, i.e., $A^2(f;n)$ is first-order unbiased for σ^2. Examples of weighting functions yielding first-order unbiased estimators for σ^2 are $f_2(t) \equiv \sqrt{840}(3t^2 - 3t + 1/2)$ and $f(t) = \sqrt{8}\pi i \cos(2\pi i t)$, $i = 1, 2, \ldots$ (see [8]). □

3 The Weighted Cramér-von Mises Estimator

In this section, we propose several estimators for σ^2 based on different Brownian bridge functionals. To parallel the discussion of §2, we define

$$C(g;n) \equiv \frac{\sum_{k=1}^{n} g(\frac{k}{n})(\sigma T_n(\frac{k}{n}))^2}{n}$$

and

$$C(g) \equiv \int_0^1 g(t)(\sigma \mathcal{B}(t))^2\, dt,$$

where $g(t)$ is a weighting function normalized so that $\mathsf{E}[C(g)] = \sigma^2$.

In the sequel, we will require a number of assumptions to hold.

Assumptions

1. The process $Y_1, Y_2, \ldots$ is stationary.

2. The constants μ and σ^2 satisfy $X_n \Rightarrow \sigma\mathcal{W}$, where

$$X_n(t) \equiv \frac{\lfloor nt \rfloor(\overline{Y}_{\lfloor nt \rfloor} - \mu)}{\sqrt{n}}.$$

3. $\sum_{k=-\infty}^{\infty} R_k = \sigma^2 > 0$.

4. $\sum_{k=1}^{\infty} k^2 |R_k| < \infty$.

5. $g''(t)$ is continuous and bounded on $[0,1]$.

6. $\int_0^1 g(t)t(1-t)\, dt = 1$ (normalizing assumption).

Remark 1 Assumptions 1–4 are conditions on the underlying stochastic processes. Glynn and Iglehart [9] list various sets of sufficient conditions for Assumption 2 to hold; these conditions usually involve moment and mixing conditions. Assumptions 3 and 4 hold for a wide variety of stochastic processes. Assumptions 5 and 6 are simply conditions on the weighting function. □

Under the Assumptions, the continuous mapping theorem implies that $C(g;n) \xrightarrow{\mathcal{D}} C(g)$. Notice that the limiting functional $C(g)$ is the weighted area under the square of a Brownian bridge; by way of contrast, the weighted area estimator's limiting functional $A^2(f)$ is the square of the weighted area under a Brownian bridge.

The distribution of $C(g)$ with constant weighting function $g_0(t) \equiv 6$, for all $t \in [0,1]$, was given by Anderson and Darling [1] and Smirnov [22]. Over sixty years ago, Cramér [4] and von Mises [23] studied statistics nearly of the form of $C(g_0;n)$ for the special case of independent and identically distributed (i.i.d.) $Y_1, Y_2, \ldots$. Anderson and Darling [1] examined the distribution of $C(g)$ with weighting function $g_{AD}(t) \equiv [t(1-t)]^{-1}$ (which does not quite meet our continuity assumption). However, the distribution of $C(g)$ with an arbitrary weighting function has not been explicitly determined; see Durbin [6] for additional details. With this historical perspective in mind, we call $C(g;n)$ the *weighted Cramér-von Mises* (CvM) estimator for σ^2.

If we observe that

$$C(g;n) = \frac{1}{n^2}\sum_{k=1}^{n} g(\frac{k}{n})k^2(\overline{Y}_n^2 - 2\overline{Y}_n\overline{Y}_k + \overline{Y}_k^2),$$

then we can give an easy $O(n)$ algorithm to calculate $C(g;n)$:

$$\begin{aligned}
&Z, S_1, S_2 \leftarrow 0\\
&\texttt{FOR } k = 1 \texttt{ TO } n\\
&\quad Z \leftarrow Z + Y_k\\
&\quad S_1 \leftarrow S_1 + g(\tfrac{k}{n})kZ\\
&\quad S_2 \leftarrow S_2 + g(\tfrac{k}{n})Z^2\\
&Z \leftarrow Z/n\\
&C(g;n) \leftarrow Z^2 \textstyle\sum_{k=1}^{n} g(\tfrac{k}{n})k^2 - 2ZS_1 + S_2
\end{aligned}$$

For now, we will be interested in moments of the CvM estimator. Our main theorem expresses the expected value of the CvM estimator $C(g;n)$ in terms of its weighting function $g(t)$ and the covariance function R_k. In what follows, we define $G \equiv \int_0^1 g(t)\,dt$.

Theorem 2 Under Assumptions 1 through 6,

$$\mathsf{E}[C(g;n)] = \sigma^2 + \frac{\gamma}{n}(G-1) + o(1/n).$$

Proof See the appendix. □

Assumptions 3 and 4 allow us to derive a useful relationship between σ^2 and σ_n^2 (cf. Schmeiser and Song [19]):

$$\begin{aligned}
\sigma_n^2 &= n\,\mathsf{Var}(\overline{Y}_n) \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathsf{Cov}(Y_i, Y_j) \\
&= R_0 + 2\sum_{i=1}^{n-1}(1-\frac{i}{n})R_i \\
&= R_0 + 2\sum_{i=1}^{\infty}(1-\frac{i}{n})R_i - 2\sum_{i=n}^{\infty}(1-\frac{i}{n})R_i \\
&= \sum_{i=-\infty}^{\infty} R_i - \frac{2}{n}\sum_{i=1}^{\infty} iR_i - 2\sum_{i=n}^{\infty}(1-\frac{i}{n})R_i \\
&= \sigma^2 + \frac{\gamma}{n} + o(1/n). \qquad (2)
\end{aligned}$$

A corollary of the main theorem that gives the analogous expression for the bias of $C(g;n)$ as an estimator for σ_n^2 follows immediately from Equation (2).

Corollary 1 Under Assumptions 1 through 6,

$$\mathsf{E}[C(g;n)] = \sigma_n^2 + \frac{\gamma}{n}(G-2) + o(1/n).$$

Some examples illustrate the consequences of Theorem 2 and Corollary 1. The simplest example assumes a constant weighting function.

Example 3 Theorem 2 implies that the CvM estimator with constant weighting function $g_0(t) = 6$ has expected value $\mathsf{E}[C(g_0;n)] = \sigma^2 + 5\gamma/n + o(1/n) = \sigma_n^2 + 4\gamma/n + o(1/n)$. □

If $G = 1$ (subject to the constraints of Assumptions 5 and 6), Theorem 2 implies that the bias of $C(g;n)$ as an estimator of σ^2 is $o(1/n)$. In this case, $C(g;n)$ is *first-order unbiased* for σ^2. Indeed, it is possible to give such a weighting.

Example 4 Consider the quadratic weighting function $g_{2;c}(t) \equiv 51 - c/2 + ct - 150t^2$, where $t \in [0,1]$ and c is real. Theorem 2 implies that $\mathsf{E}[C(g_{2;c};n)] = \sigma^2 + o(1/n)$. □

Similarly, if $G = 2$ (subject to the constraints of Assumptions 5 and 6), Corollary 1 implies that $C(g;n)$ is first-order unbiased for σ_n^2.

Example 5 Consider the quadratic weighting function $\hat{g}_{2;c}(t) \equiv 42 - c/2 + ct - 120t^2$, where $t \in [0,1]$ and c is real. Since $G = \int_0^1 \hat{g}_{2;c}(t)\,dt = 2$, Corollary 1 implies that $\mathsf{E}[C(\hat{g}_{2;c}; n)] = \sigma_n^2 + o(1/n)$. □

The choice of weighting function $g(t)$ clearly affects the variances of $C(g;n)$ and $C(g)$. (The choice of weighting function $f(t)$ affects the variance of $A^2(f;n)$, but it does *not* affect the variance of the area estimator's limiting functional, $A^2(f)$, in which case $\mathsf{Var}(A^2(f)) = 2\sigma^4$.) To see how, we shall calculate $\mathsf{Var}(C(g))$. First, we have

Lemma 1 (Patel and Read [16, p. 309]). If R and S are jointly normal, then $\mathsf{Cov}(R^2, S^2) = 2\mathsf{Cov}^2(R,S)$.

This immediately yields the following theorem on the limiting variance of $C(g;n)$.

Theorem 3 In addition to Assumptions 1 through 6, suppose that the $C^2(g;n)$'s ($n = 1, 2, \ldots$) are uniformly integrable. Then $\mathsf{Var}(C(g;n)) \to \mathsf{Var}(C(g))$, where

$$\begin{aligned}
\mathsf{Var}(C(g)) &= \sigma^4 \int_0^1 \int_0^1 g(s)g(t)\mathsf{Cov}(\mathcal{B}^2(s), \mathcal{B}^2(t))\,ds\,dt \\
&= 2\sigma^4 \int_0^1 \int_0^1 g(s)g(t)\mathsf{Cov}^2(\mathcal{B}(s), \mathcal{B}(t))\,ds\,dt \\
&\qquad \text{(by Lemma 1)} \\
&= 4\sigma^4 \int_0^1 g(t)(1-t)^2 \int_0^t g(s)s^2\,ds\,dt.
\end{aligned}$$

Example 6 Consider the constant weighting function $g_0(t) = 6$ from Example 3. Theorem 3 implies that $\mathsf{Var}(C(g_0)) = 4\sigma^4/5$. This limiting variance is significantly smaller than that of the area estimator, for which $\mathsf{Var}(A(f)) = 2\sigma^4$ (Theorem 1). □

Example 7 Consider the quadratic weighting function $g_{2;c}(t)$ from Example 4; this weighting function yields a first-order unbiased estimator for σ^2. Theorem 3 gives $\mathsf{Var}(C(g_{2;c})) = (c^2 - 300c + 26856)\sigma^4/2520$, a quantity that is minimized by the weighting function $g_2^\star(t) \equiv g_{2;150}(t)$, whence $\mathsf{Var}(C(g_2^\star)) = 121\sigma^4/70$. This limiting variance is larger than that of the CvM estimator using the constant weighting function $g_0(t)$ (Example 6); of course, the estimator for σ^2 based on $g_0(t)$ is somewhat biased (Example 3). □

Example 8 For completeness, consider the quadratic weighting function $\hat{g}_{2;c}(t)$ from Example 5; this weighting function yields a first-order unbiased estimator for σ_n^2. Theorem 3 gives $\mathsf{Var}(C(\hat{g}_{2;c})) = (c^2 - 240c + 18144)\sigma^4/2520$. This variance is minimized by the weighting function $\hat{g}_2^\star(t) \equiv \hat{g}_{2;120}(t)$, in which case $\mathsf{Var}(C(\hat{g}_2^\star)) = 52\sigma^4/35$. Although this limiting variance is lower than that of the CvM estimator using $g_2^\star(t)$, the minimum-variance first-order unbiased quadratic weighting function (Example 7), it is still larger than that of the unweighted estimator using $g_0(t)$ (Example 6). □

Ideally, we would like to choose a weighting function that minimizes the variance of the CvM estimator for σ^2 while satisfying the first-order unbiasedness and normalizing constraints; i.e., find $g(t)$ that minimizes $\mathsf{Var}(C(g))$ subject to

$$G = 1 = \int_0^1 g(t)t(1-t)\,dt. \tag{3}$$

With this goal in mind, suppose that $g(t)$ can be written as an m-degree polynomial in t, i.e.,

$$g_m(t) \equiv \sum_{i=0}^{m} c_i t^i, \quad t \in [0,1],$$

for coefficients $c_0, c_1, \ldots, c_m$ and fixed m. After some algebra, the problem becomes that of finding the coefficients that minimize

$$\mathsf{Var}(C(g_m)) = 8\sum_{i=0}^{m}\sum_{j=0}^{m} \frac{c_i c_j}{(j+3)\prod_{k=4}^{6}(i+j+k)}$$

subject to

$$\sum_{i=0}^{m} \frac{c_i}{i+1} = 1 \quad \text{and} \quad \sum_{i=0}^{m} \frac{c_i}{(i+2)(i+3)} = 1.$$

We can use Lagrangian multipliers to solve the above system. Here the Lagrangian is given by

$$\begin{aligned}
&\mathcal{L}(c_0, c_1, \ldots, c_m; \lambda_1, \lambda_2) \\
&= 8\sum_{i=0}^{m}\sum_{j=0}^{m} \frac{c_i c_j}{(j+3)\prod_{k=4}^{6}(i+j+k)} - \lambda_1\left(\sum_{i=0}^{m}\frac{c_i}{(i+2)(i+3)} - 1\right) - \lambda_2\left(\sum_{i=0}^{m}\frac{c_i}{i+1} - 1\right),
\end{aligned}$$

where λ_1 and λ_2 are constants. One takes the $m+3$ partial derivatives of $\mathcal{L}$, sets the resulting equations to zero, and solves the resulting system of linear equations for the $m+3$ unknown coefficients.

Example 9 It is easy to show via the Lagrangian method that the optimal-variance, first-order unbiased, quadratic *and* cubic polynomial weighting function is $g_2^\star(t) = -24 + 150t - 150t^2$, the choice studied in Example 7. The best quartic turns out to be

$$g_4^\star(t) \equiv \frac{-1310}{21} + \frac{19270t}{21} - \frac{25230t^2}{7} + \frac{16120t^3}{3} - \frac{8060t^4}{3},$$

in which case $\mathsf{Var}(C(g_4^\star)) = 1.042\sigma^4$. We can go further. For example, the polynomial weighting function of degree $m = 6$ that minimizes $\mathsf{Var}(C(g_6))$ subject to the constraints (3) is given by $g_6^\star(t) \equiv \sum_{i=0}^{6} c_i t^i$, where the c_i's are as follows.

i	0	1	2	3	4	5	6
c_i	-132.9358	3439.9542	-26622.7987	93037.7083	-163198.9022	140016.0576	-46672.0191

This choice of weights yields the optimal $\mathsf{Var}(C(g_6^*)) = 0.8093\sigma^4$, which is comparable to the variance of the unweighted (albeit biased) estimator $C(g_0; n)$. □

Remark 2 In order to achieve further variance savings, we can continue to increase the degree of the polynomial weighting function. However, the magnitudes of the resulting coefficients become quite large, and one must be careful to avoid round-off error as well as deleterious second-order effects for small sample sizes. □

4 Some Analytical Examples

This section presents exact analytical results involving specific stochastic processes. We shall first obtain some useful expressions for the expected values and variances of the area and CvM estimators. We assume in the sequel that Assumptions 1 through 6 are still in effect.

We begin with an intermediate result on the area estimator.

$$\begin{aligned} \mathsf{E}[A^2(f;n)] &= \mathsf{Var}(A(f;n)) \\ &= \frac{\sigma^2}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} f(\frac{i}{n})f(\frac{j}{n})\mathsf{Cov}(T_n(\frac{i}{n}), T_n(\frac{j}{n})) \\ &= \frac{1}{n^3}\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} f(\frac{i}{n})f(\frac{j}{n})ij\mathsf{Cov}(\overline{Y}_n - \overline{Y}_i, \overline{Y}_n - \overline{Y}_j). \end{aligned} \tag{4}$$

Further, if $A(f;n)$ is normal, then Lemma 1 implies that

$$\mathsf{Var}(A^2(f;n)) = 2\mathsf{Var}^2(A(f;n)) = 2(\mathsf{E}[A^2(f;n)])^2. \tag{5}$$

The analogous result on the expected value of the CvM estimator is derived next.

$$\begin{aligned} \mathsf{E}[C(g;n)] &= \frac{\sigma^2}{n}\sum_{k=1}^{n} g(\frac{k}{n})\mathsf{E}[T_n^2(\frac{k}{n})] \\ &= \frac{\sigma^2}{n}\sum_{k=1}^{n-1} g(\frac{k}{n})\mathsf{Var}(T_n(\frac{k}{n})) \\ &= \frac{1}{n^2}\sum_{k=1}^{n-1} g(\frac{k}{n})k^2\mathsf{Var}(\overline{Y}_n - \overline{Y}_k). \end{aligned} \tag{6}$$

In addition to the standing Assumptions, suppose that $Y_1, Y_2, \ldots, Y_n$ are jointly normal. Then

$$\begin{aligned}
\mathsf{Var}(C(g;n)) &= \frac{\sigma^4}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} g(\frac{i}{n})g(\frac{j}{n})\mathsf{Cov}(T_n^2(\frac{i}{n}), T_n^2(\frac{j}{n})) \\
&= \frac{2\sigma^4}{n^2}\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} g(\frac{i}{n})g(\frac{j}{n})\mathsf{Cov}^2(T_n(\frac{i}{n}), T_n(\frac{j}{n})) \\
&\quad \text{(by Lemma 1)} \\
&= \frac{2}{n^4}\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} g(\frac{i}{n})g(\frac{j}{n}) i^2 j^2 \mathsf{Cov}^2(\overline{Y}_n - \overline{Y}_i, \overline{Y}_n - \overline{Y}_j). \qquad (7)
\end{aligned}$$

We now have at our disposal the machinery to study specific examples in which we calculate the exact expected values and variances of $A^2(f;n)$ and $C(g;n)$ for various weighting functions.

In particular, for the remainder of this section, we shall work with a first-order moving average [MA(1)] process, $Y_{i+1} = \theta\epsilon_i + \epsilon_{i+1}$, $i = 1, 2, \ldots$, where the ϵ_i's are i.i.d. $\mathsf{Nor}(0,1)$; so $R_0 = 1 + \theta^2$, $R_{\pm 1} = \theta$, and $R_k = 0$, otherwise. One can derive

$$\mathsf{Var}(\overline{Y}_j) = \frac{(1+\theta)^2}{j} - \frac{2\theta}{j^2} = \frac{\sigma^2}{j} + \frac{\gamma}{j^2} \qquad (8)$$

and

$$\mathsf{Cov}(\overline{Y}_j, \overline{Y}_k) = \frac{(1+\theta)^2}{j} - \frac{\theta}{jk} = \frac{\sigma^2}{j} + \frac{\gamma}{2jk} \quad \text{for } k < j, \qquad (9)$$

where $\sigma^2 = (1+\theta)^2$ and $\gamma = -2\theta$. Note that for the MA(1) process, Equation (8) implies that $\sigma_n^2 = \sigma^2 + \frac{\gamma}{n}$, a result that agrees with Equation (2).

Example 10 We concentrate here on some area estimator expectation results for the MA(1) process. For the constant weighting function $f_0(t) = \sqrt{12}$, Equations (4), (8), and (9) show that

$$\mathsf{E}[A^2(f_0;n)] = (\sigma^2 + \frac{3\gamma}{n})(1 - \frac{1}{n^2}) = \sigma^2 + \frac{3\gamma}{n} + o(1/n),$$

as implied by Example 1. For the first-order unbiased weighting function $f_2(t) = \sqrt{840}(3t^2 - 3t + 1/2)$, we have (also see [8])

$$\begin{aligned}
\mathsf{E}[A^2(f_2;n)] &= \frac{\sigma^2(2n^6 + 7n^4 + 63n^2 - 72) + 21\gamma(2n^4 - 5n^3 + 10n^2 + 5n - 12)}{2n^6} \\
&= \sigma^2 + \frac{7(\sigma^2 + 6\gamma)}{2n^2} + O(n^{-3}) \\
&= \sigma^2 + o(1/n),
\end{aligned}$$

which is indeed first-order unbiased for σ^2 and thus is in accord with Example 2. □

Example 11 We consider CvM expectation results for the MA(1) process. For the constant weighting function $g_0(t) = 6$, Equations (6), (8), and (9) show that

$$\begin{aligned} \mathsf{E}[C(g_0; n)] &= \sigma^2 + \frac{5\gamma}{n} - \frac{\sigma^2 + 6\gamma}{n^2} + \frac{\gamma}{n^3} \\ &= \sigma^2 + \frac{5\gamma}{n} + o(1/n), \end{aligned}$$

as implied by Example 3. For the quadratic weighting function $g_{2;c}(t) = 51 - c/2 + ct - 150t^2$, we get

$$\begin{aligned} \mathsf{E}[C(g_{2;c}; n)] &= \frac{\sigma^2(n^4 + 4n^2 - 5)}{n^4} + \frac{\gamma(24n^3 - 29n^2 + 5)}{n^5} \\ &= \sigma^2 + \frac{4(\sigma^2 + 6\gamma)}{n^2} + O(n^{-3}) \\ &= \sigma^2 + o(1/n). \end{aligned}$$

This demonstrates that $g_{2;c}(t)$ yields a first-order unbiased estimator for σ^2 (independent of the choice of c), as implied by Example 4. For the optimal quartic weighting function $g_4^*(t)$, we have (after a great deal of algebra)

$$\mathsf{E}[C(g_4^*; n)] = \sigma^2 + \frac{655(\sigma^2 + 6\gamma)}{63n^2} + O(n^{-4}) = \sigma^2 + o(1/n),$$

which shows that this weighting function yields a first-order unbiased estimator for σ^2, as anticipated by Example 9. Similarly, the optimal sixth-degree weighting function $g_6^*(t)$ gives

$$\mathsf{E}[C(g_6^*; n)] = \sigma^2 + \frac{22.156(\sigma^2 + 6\gamma)}{n^2} + o(n^{-2}) = \sigma^2 + o(1/n),$$

so that this weighting function also produces a first-order unbiased estimator for σ^2. Notice that the n^{-2} term in the expression for $\mathsf{E}[C(g_6^*; n)]$ is quite large compared to the n^{-2} terms from the estimators incorporating quadratic or quartic weights; in fact, as alluded to by Remark 2, large values of n are required before the second-order term in $\mathsf{E}[C(g_6^*; n)]$ becomes insignificant.

For completeness, we give results on the quadratic weighting function $\hat{g}_{2;c}(t) = 42 - c/2 + ct - 120t^2$. In this case, we find that (after some algebra)

$$\begin{aligned} \mathsf{E}[C(\hat{g}_{2;c}; n)] &= \frac{\sigma^2(n^4 + 3n^2 - 4)}{n^4} + \frac{\gamma(n^4 + 18n^3 - 23n^2 + 4)}{n^5} \\ &= \sigma^2 + \frac{\gamma}{n} + \frac{3(\sigma^2 + 6\gamma)}{n^2} + O(n^{-3}) \\ &= \sigma_n^2 + o(1/n), \end{aligned}$$

which demonstrates that this weighting function produces a first-order unbiased estimator for σ_n^2 (independent of the choice of c), as implied by Example 5. □

Example 12 Since the MA(1) is a jointly normal process, we can easily derive area estimator variance results for it. Equation (5) and Example 10 imply that for weighting functions $f_0(t) = \sqrt{12}$ and $f_2(t) = \sqrt{840}(3t^2 - 3t + 1/2)$, we have $\mathsf{Var}(A^2(f;n)) = 2\sigma^4 + o(1)$. These results make sense in light of Theorem 1. □

Example 13 We examine the variance of the CvM estimator for various weighting functions. For the constant weighting function $g_0(t) = 6$, Equation (7) gives us (after tedious but straightforward algebra)

$$\mathsf{Var}(C(g_0;n)) = \frac{4\sigma^4}{5} + \frac{16\gamma\sigma^2}{5n} + O(n^{-2}) = \frac{4\sigma^4}{5} + o(1),$$

as implied by Example 6. For the (variance-optimal and first-order unbiased for σ^2) quadratic weighting function $g_2^\star(t) = -24 + 150t - 150t^2$, some algebra yields

$$\mathsf{Var}(C(g_2^\star;n)) = 1.7286\sigma^4 + \frac{4.0571\gamma\sigma^2}{n} + O(n^{-2}) = 1.7286\sigma^4 + o(1),$$

as implied by Examples 7 and 9. For the (variance-optimal and first-order unbiased for σ^2) quartic weighting function $g_4^\star(t)$, we can obtain

$$\mathsf{Var}(C(g_4^\star;n)) = 1.0418\sigma^4 + \frac{3.8235\gamma\sigma^2}{n} + O(n^{-2}) = 1.0418\sigma^4 + o(1),$$

as implied by Example 9.

Finally, for the (variance-optimal and first-order unbiased for σ_n^2) quadratic weighting function $\hat{g}_2^\star(t) = -18 + 120t - 120t^2$, some algebra yields

$$\mathsf{Var}(C(\hat{g}_2^\star;n)) = 1.4857\sigma^4 + \frac{3.6571\gamma\sigma^2}{n} + O(n^{-2}) = 1.4857\sigma^4 + o(1),$$

as implied by Example 8. □

We see from the above examples that the area and CvM estimators behave as advertised on the simple analytical MA(1) example. The CvM estimator using the Anderson-Darling weighting function (which *fails* to satisfy some of the Assumptions) does not behave so nicely.

Example 14 To complete our series of examples with the MA(1) process, we consider the expected value of the Anderson-Darling estimator, i.e., the CvM estimator with weighting function $g_{AD}(t) = [t(1-t)]^{-1}$. Then it can be shown that

$$\begin{aligned}\mathsf{E}[C(g_{AD};n)] &= \sigma^2(1-\frac{1}{n}) - \frac{\gamma}{n}\left(1 - \frac{1}{n} - 2\sum_{k=1}^{n-1}\frac{1}{k}\right)\\ &\approx \sigma^2(1-\frac{1}{n}) - \frac{\gamma}{n}(1 - \frac{1}{n} - 2(\ell\mathrm{n}(n-1) + c_e))\\ &= \sigma^2 + \frac{2\gamma\,\ell\mathrm{n}(n-1)}{n} + o\left(\frac{\ell\mathrm{n}(n)}{n}\right),\end{aligned}$$

where $c_e \approx 0.577216$ is Euler's constant. Although this estimator is asymptotically unbiased for σ^2, the convergence rate of the expectation to σ^2 is *much slower* than those of the previous examples. $\square$

We resort to Monte Carlo simulation in the next section to empirically evaluate the performance characteristics of the various estimators on more complicated stochastic processes.

5 Empirical Examples

In this section, we present empirical examples illustrating the performance characteristics of the following variance estimators:

- $A^2(f_0;n)$ — unweighted area estimator.
- $A^2(f_2;n)$ — first-order unbiased quadratic area estimator for σ^2.
- $C(g_0;n)$ — unweighted CvM estimator.
- $C(g_{AD};n)$ — Anderson-Darling estimator.
- $C(g_2^\star;n)$ — minimum-variance first-order unbiased quadratic CvM estimator for σ^2.
- $C(g_4^\star;n)$ — minimum-variance first-order unbiased quartic CvM estimator for σ^2.
- $C(g_6^\star;n)$ — minimum-variance first-order unbiased sixth-degree CvM estimator for σ^2.
- $C(\hat{g}_2^\star;n)$ — minimum-variance first-order unbiased quadratic CvM estimator for σ_n^2.

These examples involve the Monte Carlo simulation of a number of stationary stochastic processes:

- The first-order autoregressive process [AR(1)], $Y_{i+1} = \phi Y_i + \epsilon_{i+1}$, $i = 1, 2, \ldots$, where the ϵ_i's are i.i.d. $\mathrm{Nor}(0, 1-\phi^2)$ with $-1 < \phi < 1$.
- The first-order exponential autoregressive process [EAR(1)], $Y_{i+1} = \phi Y_i + \epsilon_{i+1}$, $i = 1, 2, \ldots$, where the ϵ_i's are i.i.d. exponential(1) with probability $1-\phi$ and 0 otherwise, and where $0 \le \phi < 1$. (See Lewis [14] for more details.)
- The M/M/1 queueing system's waiting-time process.

Table 1: Estimated Expected Values of Various Variance Estimators — AR(1), $\phi = 0.9$. (Note that $\sigma^2 = 19.0$ for this process.)

n	$A^2(f_0;n)$	$A^2(f_2;n)$	$C(g_0;n)$	$C(g_{AD};n)$	$C(g_2^*;n)$	$C(g_4^*;n)$	$C(g_6^*;n)$	$C(\hat{g}_2^*;n)$	σ_n^2
4	0.289	0.402	0.225	0.177	0.316	0.294	0.222	0.298	3.5245
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	
8	0.974	0.966	0.698	0.576	0.940	0.757	0.674	0.892	6.1855
	(0.004)	(0.004)	(0.003)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)	
16	2.837	2.677	2.041	1.701	2.747	2.168	1.847	2.606	9.8346
	(0.013)	(0.012)	(0.008)	(0.007)	(0.012)	(0.008)	(0.007)	(0.011)	
32	6.533	6.439	4.934	4.141	6.580	5.262	4.481	6.251	13.5681
	(0.029)	(0.029)	(0.019)	(0.015)	(0.028)	(0.019)	(0.016)	(0.026)	
64	11.269	11.889	9.186	7.812	11.925	9.903	8.645	11.377	16.1908
	(0.050)	(0.053)	(0.033)	(0.026)	(0.050)	(0.034)	(0.028)	(0.046)	
128	14.914	16.200	13.148	11.476	16.119	14.395	12.811	15.525	17.5938
	(0.067)	(0.072)	(0.044)	(0.035)	(0.066)	(0.047)	(0.039)	(0.062)	
256	16.836	18.017	15.752	14.236	17.968	17.207	16.073	17.525	18.2969
	(0.076)	(0.081)	(0.049)	(0.040)	(0.074)	(0.055)	(0.046)	(0.068)	
512	17.989	18.824	17.349	16.164	18.783	18.507	17.924	18.496	18.6484
	(0.081)	(0.085)	(0.052)	(0.043)	(0.078)	(0.059)	(0.051)	(0.072)	
1024	18.372	18.875	18.084	17.285	18.838	18.867	18.617	18.687	18.8242
	(0.082)	(0.084)	(0.053)	(0.044)	(0.078)	(0.060)	(0.052)	(0.072)	
2048	18.799	19.064	18.599	18.073	19.027	19.091	18.932	18.941	18.9121
	(0.084)	(0.085)	(0.053)	(0.045)	(0.079)	(0.061)	(0.054)	(0.073)	

For both the AR(1) and EAR(1) processes, the covariance function $R_k = \phi^{|k|}$, $k = 0, \pm 1, \pm 2, \ldots$. The covariance function of the M/M/1 waiting time process is more complicated (cf. Daley [5]).

We simulated the above stochastic processes over a variety of parameter values; representative results are presented in Table 1 (AR(1) with $\phi = 0.9$), Table 2 (EAR(1) with $\phi = 0.9$), and Table 3 (M/M/1 waiting time process with arrival rate 0.8 and service rate 1.0). Each table entry in a row is based on the same 100,000 independent replications of the stochastic process. The number in parentheses below an entry is the standard error of that entry. Each of the replications was initialized from the appropriate *steady-state* distribution. All uniform [normal] random variates were generated from algorithm `UNIF` [`TRPNRM`] in Bratley, Fox, and Schrage [3]; exponential deviates used inversion; the M/M/1 waiting-time process was generated from an algorithm due to Schmeiser [18].

Table 2: Estimated Expected Values of Various Variance Estimators — EAR(1), $\phi = 0.9$. (Note that $\sigma^2 = 19.0$ for this process.)

n	$A^2(f_0;n)$	$A^2(f_2;n)$	$C(g_0;n)$	$C(g_{AD};n)$	$C(g_2^\star;n)$	$C(g_4^\star;n)$	$C(g_6^\star;n)$	$C(\hat{g}_2^\star;n)$	σ_n^2
4	0.295	0.410	0.229	0.180	0.322	0.299	0.226	0.303	3.5245
	(0.004)	(0.006)	(0.003)	(0.002)	(0.005)	(0.004)	(0.003)	(0.004)	
8	0.979	0.971	0.703	0.580	0.946	0.764	0.677	0.897	6.1855
	(0.009)	(0.010)	(0.007)	(0.005)	(0.010)	(0.007)	(0.006)	(0.009)	
16	2.859	2.700	2.059	1.717	2.770	2.188	1.866	2.627	9.8346
	(0.021)	(0.021)	(0.014)	(0.011)	(0.021)	(0.015)	(0.012)	(0.019)	
32	6.554	6.504	4.960	4.160	6.630	5.281	4.497	6.296	13.5681
	(0.040)	(0.041)	(0.027)	(0.022)	(0.040)	(0.029)	(0.024)	(0.037)	
64	11.303	11.892	9.201	7.824	11.934	9.932	8.570	11.388	16.1908
	(0.063)	(0.067)	(0.043)	(0.035)	(0.062)	(0.047)	(0.039)	(0.058)	
128	15.008	16.241	13.218	11.536	16.198	14.491	12.884	15.602	17.5938
	(0.078)	(0.084)	(0.053)	(0.044)	(0.077)	(0.059)	(0.050)	(0.072)	
256	16.942	18.125	15.821	14.289	18.081	17.246	16.138	17.629	18.2969
	(0.082)	(0.088)	(0.055)	(0.047)	(0.080)	(0.063)	(0.055)	(0.074)	
512	17.984	18.810	17.345	16.161	18.771	18.511	17.925	18.486	18.6484
	(0.085)	(0.087)	(0.056)	(0.048)	(0.080)	(0.064)	(0.057)	(0.074)	
1024	18.374	18.841	18.088	17.286	18.835	18.925	18.592	18.685	18.8242
	(0.084)	(0.086)	(0.055)	(0.047)	(0.079)	(0.063)	(0.055)	(0.073)	
2048	18.719	18.955	18.552	18.029	18.985	19.039	18.849	18.898	18.9121
	(0.085)	(0.086)	(0.054)	(0.046)	(0.079)	(0.062)	(0.055)	(0.073)	

Table 3: Estimated Expected Values of Various Variance Estimators — M/M/1 Waiting Time Process; arrival rate = 0.8, service rate = 1.0

n	$A^2(f_0;n)$	$A^2(f_2;n)$	$C(g_0;n)$	$C(g_{AD};n)$	$C(g_2^\star;n)$	$C(g_4^\star;n)$	$C(g_6^\star;n)$	$C(\hat{g}_2^\star;n)$
16	30.9	28.0	21.6	18.0	29.1	22.9	19.5	27.6
	(0.2)	(0.2)	(0.1)	(0.1)	(0.2)	(0.1)	(0.1)	(0.2)
32	93.9	85.3	66.4	55.5	89.5	70.5	59.7	85.1
	(0.6)	(0.5)	(0.4)	(0.3)	(0.5)	(0.4)	(0.3)	(0.5)
64	251.7	235.7	182.9	153.2	245.5	194.4	165.2	235.0
	(1.7)	(1.6)	(1.1)	(0.9)	(1.6)	(1.1)	(0.9)	(1.5)
128	574.0	562.2	434.5	365.7	577.1	464.1	396.1	542.6
	(4.5)	(4.4)	(3.0)	(2.4)	(4.4)	(3.1)	(2.5)	(4.0)
256	1002.1	1035.6	810.4	690.7	1047.3	875.5	756.2	1019.0
	(9.1)	(9.2)	(6.2)	(5.0)	(9.0)	(6.5)	(5.3)	(8.7)
512	1415.6	1541.2	1228.0	1065.4	1532.2	1327.5	1180.1	1493.9
	(14.4)	(15.4)	(10.1)	(8.3)	(14.6)	(10.9)	(9.0)	(14.6)
1024	1708.8	1827.8	1563.1	1398.0	1824.8	1710.9	1569.0	1769.4
	(18.4)	(19.3)	(13.2)	(11.1)	(18.0)	(15.2)	(12.8)	(16.3)
2048	1831.7	1919.2	1749.5	1616.4	1922.7	1879.0	1803.4	1921.2
	(17.0)	(17.5)	(12.6)	(11.2)	(15.8)	(15.0)	(14.0)	(15.7)
4096	1914.4	1986.6	1867.3	1767.7	1979.7	1958.5	1918.1	1954.2
	(14.2)	(14.7)	(10.6)	(9.9)	(13.0)	(12.2)	(12.3)	(13.2)
8192	1970.7	2016.0	1940.9	1870.6	2008.2	2009.8	1980.5	1944.3
	(12.3)	(12.4)	(8.9)	(8.5)	(11.1)	(10.3)	(10.2)	(10.3)

6 Discussion

This section summarizes and discusses the exact and estimated expectation and variance results for the variance estimators examined in §§2–5. Recall that we obtained exact results for area estimators in §2 and for CvM estimators in §3. We also gave exact results for a specific stochastic process, the MA(1), in §4 and empirical results for the AR(1), EAR(1), and M/M/1 in §5. The estimated expected values given in Tables 1–3 are based on 100,000 replications; thus, one can obtain estimated variance results from the Monte Carlo runs by squaring the estimated standard errors (in parentheses below the estimated expected values) and then multiplying by 100,000.

For each of the stochastic processes under study, the expected value of the unweighted area estimator $A^2(f_0;n)$ converged relatively slowly to σ^2 as n increased. This phenomenon is due to the estimator's comparatively high $\mathsf{Bias}(A^2(f_0;n)) \approx 3\gamma/n$ (Example 1; also see [17]).

The expected value of the unweighted CvM estimator $C(g_0;n)$ converged more slowly than that of $A^2(f_0;n)$ to σ^2. This makes sense since $\mathsf{Bias}(C(g_0;n)) \approx 5\gamma/n$ (Example 3) is higher than $A^2(f_0;n)$'s. On the plus side, we see that for large n,

$$\mathsf{Var}(C(g_0;n))/\mathsf{Var}(A^2(f_0;n)) \approx \frac{0.8\sigma^4}{2\sigma^4} = 0.4,$$

as suggested by Theorem 1 and Example 6.

The expected value of the Anderson-Darling CvM estimator $C(g_{AD};n)$ converged even more slowly than that of $C(g_0;n)$ to σ^2. Although we did not prove the inferiority of the convergence rate of $\mathsf{E}[C(g_{AD};n)]$ to σ^2 for general stationary processes, the evidence provided by Example 14 and the empirical work seems to be overwhelming. Some calculus shows that

$$\mathsf{Var}(C(g_{AD})) = (\frac{\pi^2}{6} - 1)\sigma^4 \approx 0.5797\sigma^4$$

(cf. [21]); but it is of little comfort that $C(g_{AD};n)$ has the lowest variance of the estimators under study.

The expected values of the first-order unbiased quadratic area estimator for σ^2, $A^2(f_2;n)$, and the minimum-variance first-order unbiased quadratic CvM estimator for σ^2, $C(g_2^\star;n)$, converged comparatively quickly to σ^2 as n increased; the rapid convergence is a direct consequence of the first-order unbiasedness of the estimators. For large n, we see from Examples 12 and 13 and the empirical tables that

$$\mathsf{Var}(C(g_2^\star;n))/\mathsf{Var}(A^2(f_2;n)) \approx 121/140,$$

as predicted by Theorem 1 and Example 7.

The minimum-variance first-order unbiased fourth- and sixth-degree CvM estimators, $C(g_4^\star;n)$ and $C(g_6^\star;n)$, respectively, possess expected values that converge to σ^2 almost (but not quite) as quickly as those of $A^2(f_2;n)$ and $C(g_2^*;n)$. A favorable property of these higher-degree estimators is that they have reduced variances; the variance improvements are along the lines described in Example 9.

Recall that the estimator $C(\hat{g}_2^*;n)$ is first-order unbiased for σ_n^2 (Example 5). We can compare the estimated expected values from Tables 1 and 2 with the corresponding actual σ_n^2-values (given in the last column of the tables). We see that the bias of $C(\hat{g}_2^*;n)$ as an estimator for σ_n^2 is about the same as $\mathsf{Bias}(A^2(f_2;n))$ and $\mathsf{Bias}(C(g_2^*;n))$; this is particularly true for large n. In addition, $\mathsf{Var}(C(\hat{g}_2^*;n))$ is only a little smaller than $\mathsf{Var}(C(g_2^*;n))$. Thus, we do not seem to gain much by using $C(\hat{g}_2^*;n)$ to estimate σ_n^2.

The bottom line: Of the estimators studied so far, it appears that $C(g_2^*;n)$ performs the best.

7 Conclusions

In this article, we introduced a class of estimators for $\sigma^2 = \lim_{n\to\infty} n\,\mathsf{Var}(\overline{Y}_n)$ that are similar to Cramér-von Mises statistics. Using appropriate weighting functions, our CvM estimators were shown to be first-order unbiased, and asymptotic variance reductions of up to 60% (compared to the weighted area estimator) were achievable. Further, the variance-optimal weighting functions can be computed independently of the output process. Our conclusions were supported with an analytical example using the MA(1) process.

Although the estimators are all asymptotically unbiased for σ^2, finite-sample bias can be substantial. Analytical and empirical work showed that the bias of the CvM estimators converged to zero at least as fast as that of the weighted area estimators, and the CvM estimators had smaller variance. Thus, if the sample size is sufficiently large, the CvM estimators proved to be more efficient than the weighted area estimators.

As discussed in [10], it is possible to augment the basic CvM variance estimator in a number of ways.

1. One can show that the unweighted CvM and area estimators are highly correlated. This suggests that certain *linear combinations* of the area and CvM estimators will give rise to estimators having comparatively lower variance.

2. All of our work so far has assumed that we have one long batch of n observations. An alternative way of organizing the data is to break the n observations into b contiguous, nonoverlapping batches, each of size m (assume $n = bm$). This leads to another interesting problem—that of examining the consequences of batching the data and then forming CvM estimators from each batch. Intuitively, batching of the data will tend to increase estimator bias while decreasing estimator variance—of

course, one can quantify the trade-off by calculating the batched CvM estimator's mean squared error. (See Schmeiser and Song [19].)

3. We can also apply the methodology of Meketon and Schmeiser [15] in which the n observations are broken into $n-m+1$ *overlapping* batches, each of size m. Although the bias of the resulting overlapped CvM estimator is the same as that of the batched CvM estimator, the overlapped estimator's variance is much smaller.

The above problems are the subjects of ongoing research and will be the topics of a future paper.

Acknowledgments

We thank George Fishman and Bruce Schmeiser for many interesting discussions. David Goldsman's work was supported by National Science Foundation Grant No. DDM-90-12020.

Appendix

This appendix contains the proof of Theorem 2. Before proving the theorem, we state and prove a series of lemmas. First, we define the cumulative sums $Z_k \equiv \sum_{j=1}^{k} Y_j$ and the *variance time curve* $V(k) \equiv \mathrm{Var}(Z_k)$, $k = 1, 2, \ldots, n$ (see [11]). Since $g(t)$ is assumed to be continuous and bounded on $[0,1]$, we denote $M \equiv \sup_{0\le t\le 1} |g(t)| < \infty$.

Lemma 2 Under the Assumptions of §3,

$$V(n) \;=\; n\sigma^2 + \gamma - 2\sum_{i=n}^{\infty}(n-i)R_i \;=\; n\sigma^2 + \gamma + o(1).$$

Proof Follows from the arguments leading to Equation (2). □

Lemma 3 (The discrete-time version of a result given in [11].) Under the Assumptions of §3, if $n \ge k$, we have

$$2\,\mathrm{Cov}(Z_n, Z_k) \;=\; V(n) + V(k) - V(n-k).$$

Proof By stationary increments,

$$\begin{aligned} V(n) &= \mathrm{Var}(Z(n) - Z(k) + Z(k)) \\ &= \mathrm{Var}(Z(n) - Z(k)) + \mathrm{Var}(Z(k)) + 2\mathrm{Cov}(Z(n) - Z(k), Z(k)) \\ &= \mathrm{Var}(Z(n-k)) + V(k) + 2\mathrm{Cov}(Z(n), Z(k)) - 2V(k). \quad \square \end{aligned}$$

Lemma 4 Under the Assumptions of §3,

$$\left|\sum_{k=1}^{n} g(\frac{k}{n})k^2\right| \leq M\sum_{k=1}^{n} k^2 = O(n^3).$$

Lemma 5 Under the Assumptions of §3,

$$\begin{aligned}
&\left|\sum_{k=1}^{n} g(\frac{k}{n})k\left[\sum_{j=n}^{\infty}(j-n)R_j + \sum_{j=k}^{\infty}(j-k)R_j - \sum_{j=n-k}^{\infty}(j-(n-k))R_j\right]\right| \\
\leq\ & M\sum_{k=1}^{n} k\left[3\sum_{j=n}^{\infty} j|R_j| + \sum_{j=k}^{n-1} j|R_j| + \sum_{j=n-k}^{n-1} j|R_j|\right] \\
=\ & M\left[\frac{3}{2}n(n+1)\sum_{j=n}^{\infty} j|R_j| + n\sum_{k=1}^{n-1} k^2|R_k|\right] \\
=\ & o(n^2).
\end{aligned}$$

Lemma 6 Under the Assumptions of §3,

$$\begin{aligned}
&\left|\sum_{k=1}^{n} g(\frac{k}{n})\sum_{j=k}^{\infty}(j-k)R_j\right| \\
&\leq\ M\sum_{k=1}^{n}\sum_{j=k}^{\infty} j|R_j| \\
&=\ M\sum_{k=1}^{n} k^2|R_k| + Mn\sum_{j=n+1}^{\infty} j|R_j| \\
&=\ o(n).
\end{aligned}$$

Lemma 7 Under the Assumptions of §3,

$$V(n)\sum_{k=1}^{n} g(\frac{k}{n})k^2 = (n\sigma^2+\gamma+o(1))\sum_{k=1}^{n} g(\frac{k}{n})k^2 = (n\sigma^2+\gamma)\sum_{k=1}^{n} g(\frac{k}{n})k^2 + o(n^3).$$

Proof Follows from Lemmas 2 and 4. □

Lemma 8 Under the Assumptions of §3,

$$\begin{aligned}
&\sum_{k=1}^{n} g(\frac{k}{n})k\mathsf{Cov}(Z_n, Z_k) \\
&= \sum_{k=1}^{n} g(\frac{k}{n})k\left[k\sigma^2 + \frac{\gamma}{2} + \sum_{j=n}^{\infty}(j-n)R_j + \sum_{j=k}^{\infty}(j-k)R_j - \sum_{j=n-k}^{\infty}(j-(n-k))R_j\right] \\
&= \sigma^2 \sum_{k=1}^{n} g(\frac{k}{n})k^2 + \frac{\gamma}{2}\sum_{k=1}^{n} g(\frac{k}{n})k + o(n^2).
\end{aligned}$$

Proof Follows from Lemmas 2, 3, and 5. □

Lemma 9 Under the Assumptions of §3,

$$\begin{aligned}
\sum_{k=1}^{n} g(\frac{k}{n})V(k) &= \sum_{k=1}^{n} g(\frac{k}{n})\left[k\sigma^2 + \gamma + 2\sum_{j=k}^{\infty}(j-k)R_j\right] \\
&= \sigma^2 \sum_{k=1}^{n} g(\frac{k}{n})k + \gamma \sum_{k=1}^{n} g(\frac{k}{n}) + o(n).
\end{aligned}$$

Proof Follows from Lemmas 2 and 6. □

Lemma 10 (Trapezoid Rule). If $e''(t)$ is continuous and bounded $\forall t \in [0,1]$, then

$$\int_0^1 e(s)\,ds = \frac{1}{n}\sum_{k=1}^{n} e(\frac{k}{n}) + \frac{e(0)-e(1)}{2n} + o(1/n).$$

We can finally prove Theorem 2.

Proof (of Theorem 2).

$$\begin{aligned}
\mathsf{E}[W^2(n)] &= \frac{1}{n^2}\sum_{k=1}^{n} g(\frac{k}{n})k^2\mathsf{E}[(\overline{Y}_n - \overline{Y}_k)^2] \\
&= \frac{V(n)}{n^4}\sum_{k=1}^{n} g(\frac{k}{n})k^2 - \frac{2}{n^3}\sum_{k=1}^{n} g(\frac{k}{n})k\mathsf{Cov}(Z_n, Z_k) + \frac{1}{n^2}\sum_{k=1}^{n} g(\frac{k}{n})V(k) \\
&= \sigma^2\left[\frac{1}{n^2}\sum_{k=1}^{n} g(\frac{k}{n})k - \frac{1}{n^3}\sum_{k=1}^{n} g(\frac{k}{n})k^2\right] + \gamma\left[\frac{1}{n^4}\sum_{k=1}^{n} g(\frac{k}{n})k^2 - \frac{1}{n^3}\sum_{k=1}^{n} g(\frac{k}{n})k + \frac{1}{n^2}\sum_{k=1}^{n} g(\frac{k}{n})\right] + o(\frac{1}{n})
\end{aligned}$$

(by Lemmas 7, 8, and 9, and algebra)

$$= \sigma^2\int_0^1 g(t)(t-t^2)\,dt + \frac{\gamma}{n}\int_0^1 g(t)(t^2-t+1)\,dt + o(1/n)$$

(by Lemma 10).

Application of Assumption 6 completes the proof. □

Remark 3 It is quite a bit easier to prove the continuous-time version of the theorem (cf. [11]).

References

[1] T. W. ANDERSON AND D. A. DARLING (1952). "Asymptotic Theory of Certain 'Goodness of Fit' Criteria Based on Stochastic Processes," *Annals of Mathematical Statistics* **23**, 193–212.

[2] P. BILLINGSLEY (1968). *Convergence of Probability Measures*, John Wiley & Sons, New York.

[3] P. BRATLEY, B. L. FOX, AND L. E. SCHRAGE (1987). *A Guide to Simulation*, 2nd Edition, Springer-Verlag, New York.

[4] H. CRAMÉR (1928). "On the Composition of Elementary Errors. Second Paper: Statistical Applications," *Skand. Aktuartidskr.* **11**, 141–180.

[5] D. J. DALEY (1968). "The Serial Correlation Coefficients of Waiting Times in a Stationary Singer Server Queue," *J. Austr. Math. Soc.* **8**, 683–699.

[6] J. DURBIN (1973). *Distribution Theory for Tests Based on the Sample Distribution Function*, Society for Industrial and Applied Mathematics, Philadelphia.

[7] K. DZHAPARIDZE (1986). *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series*, Springer-Verlag, New York.

[8] R. D. FOLEY AND D. GOLDSMAN (1990). "Confidence Intervals Using Orthonormally Weighted Standardized Time Series," Technical Report, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia.

[9] P. GLYNN AND D. L. IGLEHART (1990). "Simulation Output Analysis Using Standardized Time Series," *Mathematics of Operations Research* **15**, 1–16.

[10] D.GOLDSMAN, K. KANG, AND A. F. SEILA (1991). "Cramér-von Mises Variance Estimators for Simulations," *Proceedings of the 1991 Winter Simulation Conference*, 916–920.

[11] D. Goldsman and M. S. Meketon (1986). "A Comparison of Several Variance Estimators," Technical Report J-85-12, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia.

[12] D. Goldsman, M. S. Meketon, and L. W. Schruben (1990). "Properties of Standardized Time Series Weighted Area Variance Estimators," *Management Science* **36**, 602–612.

[13] D. Goldsman and L. W. Schruben (1990). "New Confidence Interval Estimators Using Standardized Time Series," *Management Science* **36**, 393–397.

[14] P. A. W. Lewis (1980). "Simple Methods for Positive-Valued and Discrete-Valued Time Series with ARMA Correlation Structure," in *Multivariate Analysis—V* (ed. P. R. Krishnaiah), 151–166, North Holland, New York.

[15] M. S. Meketon and B. W. Schmeiser (1984). "Overlapping Batch Means: Something for Nothing?" *Proceedings of the 1984 Winter Simulation Conference*, 227–230.

[16] J. K. Patel and C. B. Read (1982). *Handbook of the Normal Distribution*, Marcel Dekker, New York.

[17] R. G. Sargent, K. Kang, and D. Goldsman (1992). "An Investigation of Finite-Sample Behavior of Confidence Interval Estimators," *Operations Research* **40**, 898–913.

[18] B. W. Schmeiser (1990). Personal Communication.

[19] B. W. Schmeiser and W.-M. T. Song (1989). "Optimal Mean-Squared-Error Batch Sizes," to appear in *Management Science*.

[20] L. W. Schruben (1983). "Confidence Interval Estimation Using Standardized Time Series," *Operations Research* **31**, 1090–1108.

[21] G. R. Shorack and J. A. Wellner (1986). *Empirical Processes with Applications to Statistics*, John Wiley & Sons, New York.

[22] N. V. Smirnov (1937). "On the Distribution of the von Mises ω^2-Criterion" (in Russian), *Matem Sbornik.* **5**, 973–993.

[23] R. von Mises (1931). *Wahrscheinlichkeitsrechnung*, Wein, Leipzig.

DISTRIBUTION LIST

Agency	No. of copies
Defense Technical Information Center Cameron Station Alexandria, VA 22314	2
Dudley Knox Library, Code 0142 Naval Postgraduate School Monterey, CA 93943	2
Office of Research Administration, Code 08 Naval Postgraduate School Monterey, CA 93943	1
Library, Center for Naval Analyses 4401 Ford Avenue Alexandria, VA 22302-0268	1
Department of Systems Management Library Code AS Naval Postgraduate School Monterey, CA 93943	1
Professor Keebom Kang, Code AS/Kk Department of Systems Management Naval Postgraduate School Monterey, CA 93943	20